Revisiting discrete logarithms in medium/small characteristic

Antoine Joux
UVSQ and CryptoExperts

June 20th, 2013
Discrete logarithms

- Given a multiplicative group $G$ with generator $g$
- Computing discrete logarithms is inversing $n \rightarrow g^n$
- Hard in general and used as a hard problem in cryptography
- Algorithmic viewpoint
  - Generic algorithms (for any $G$)
  - Specific algorithms (make use of group representation)
Generic algorithms: Pohlig-Hellman

- Given a multiplicative group $G$ with generator $g$
- Given $|G| = \prod_{i=1}^{k} p_i^{e_i}$
- To compute dlogs in $G$, it suffices to compute dlogs in:

$$G_i = \langle g \rangle_{G/p_i}$$  (Group of order $p_i$)
Generic algorithms: $|G| = p$

- There exist algorithms with complexity $O(\sqrt{p})$ to solve:
  $$y = g^n$$

- Baby-step giant-step (let $R = \lceil \sqrt{p} \rceil$):
  - Create list $y, y/g, \ldots, y/g^{R-1}$
  - Create list $1, h, h^2, \ldots, h^{R-1}$, where $h = g^R$
  - Find collision

- Can be improved to memoryless algorithms using cycle finding techniques
Classical groups for Dlog in Cryptography

- Integers modulo $p$
- More general finite fields $\mathbb{F}_{p^k}$
- Elliptic curves over finite fields
Index calculus algorithms

- Relation generation phase
  - Generates many sparse equations
  - Modulo group order for discrete log
    (Modulo 2 for factoring)

- Linear algebra phase
  - Large sparse system
  - Numbers of unknowns in range up to dozens of millions
  - Number of equations potentially very large
  - Need to use large computers to solve such systems

- Individual logarithm phase
Complexity of Index calculus algorithms

Write:

\[ L_Q(\beta, c) = \exp((c + o(1))(\log Q)^\beta (\log \log Q)^{1-\beta}). \]

Complexity of dlogs with index calculus algorithms

- Number field sieve (\( p \) large):
  \[ L_p \left( \frac{1}{3}, \left( \frac{64}{9} \right)^{1/3} \right) \]

- Number field sieve (\( p \) medium to large, \( Q = p^k \)):
  \[ L_Q \left( \frac{1}{3}, \left( \frac{128}{9} \right)^{1/3} \right) \]

- Function field sieve (\( p \) small to medium, \( Q = p^k \)):
  \[ L_Q \left( \frac{1}{3}, \left( \frac{32}{9} \right)^{1/3} \right), \]

the constant is reduced for some specific balance of \( p \) and \( k \).

Antoine Joux UVSQ and CryptoExperts
Revisiting discrete logarithms in medium/small characteristic
Discrete Logarithms in the Medium prime case [JL06]

- Finite field of the form $\mathbb{F}_{p^k}$

- Choose two univariate polynomials $f_1$ and $f_2$
  - with degrees $d_1$ and $d_2$ and $d_1 d_2 \geq k$.
  - Such that $x - f_1(f_2(x))$ has:
    - an irreducible factor of degree $k$ (modulo $p$).

- This defines the finite field by the relations:
  - $x = f_1(y)$ and $y = f_2(x)$
Optimal for \( p = L_{p^k}(1/3) \)

Choose smoothness basis \( x - \alpha \) and \( y - \alpha \)

Consider elements:

\[
xy + ay + bx + c = x f_2(x) + af_2(x) + bx + c \\
= y f_1(y) + ay + bf_1(y) + c
\]

When both sides split \( \Rightarrow \) Relation

Heuristic cost of finding relation (sieving):

\[
(d_1 + 1)! (d_2 + 1)!
\]

Individual log. descent negligible compared to initial phase
Nice special case – Kummer extensions

- Assume $k | p - 1$, then $\mathbb{F}_{p^k}$ can be defined by $x^k - t$
- If $k = d_1 d_2 - 1$, let $y = x^{d_1}$ and $tx = y^{d_2}$
- Reduces size of smoothness basis by $k$
  - Indeed:

$$
(X + \alpha)^p = X^p + \alpha = t^{(p-1)/k} X + \alpha = \mu(X + \alpha/\mu),
$$
$$
(Y + \alpha)^p = \mu^{d_1}(Y + \alpha/\mu^{d_1}).
$$

where $\mu$ is a $k$-th root of unity in $\mathbb{F}_p$.

- Can be generalized to $k = d_1 d_2 + 1$ using $y = x^{d_1}$ and $x = t/y^{d_2}$
Linear change of variables [J13]

- Further restrict to $y = x^{d_1}$
- Then:
  
  $$xy + ay + bx + c = x^{d_1+1} + ax^{d_1} + bx + c$$

- Perform change of variable: $x = aX$, we get:

  $$a^{d_1+1}(X^{d_1+1} + X^{d_1} + b \cdot a^{-d_1}(X + c/(ab))).$$

- Change of variable does not affect splitting property
- One good left-hand side $\Rightarrow p$ good left-hand sides
- Amortized cost of relation reduced to

  $$\left(\frac{(d_1 + 1)!}{p - 1} + 1\right) \cdot (d_2 + 1)!$$
Case of Kummer extensions

- Assume $k|p - 1$, i.e. $\mathbb{F}_{p^k}$ can be defined by $x^k - t$

- If $k = d_1 d_2 - 1$, let $y = x^{d_1}$ and $tx = y^{d_2}$
  - $x^{d_1 + 1} + ax^{d_1} + bx + c \Rightarrow a^{d_1 + 1}(X^{d_1 + 1} + X^{d_1} + b \cdot a^{-d_1}(X + c/(ab))$.
  - $(y^{d_2 + 1} + by^{d_2})/t + ay + c \Rightarrow b^{d_2 + 1}((Y^{d_2 + 1} + Y^{d_2})/t + a \cdot b^{-d_2}(Y + c/(ab))$.

- In both cases $\lambda = c/(ab)$ is shared by the two sides
Assume that:

\[ X^{d_1+1} + X^{d_1} + \theta_X(X + \lambda) \text{ splits and} \]
\[ (Y^{d_2+1} + Y^{d_2})/t + \theta_Y(Y + \lambda) \text{ splits.} \]

Find \( a \) and \( b \) such that \( \theta_X = b \cdot a^{-d_1} \) and \( \theta_Y = a \cdot b^{-d_2} \)?

This implies \( \theta_X^{d_2} \theta_Y = a^{-d_1d_2+1} = a^{-k} \).

Possible iff \( \theta_X^{d_2} \theta_Y \) is a \( k \)-th power

Gives \( k \) (conjugate) solutions!

From \( a \) recover \( b \) and \( c \)

Roots obtained by change of variable
Impact in the medium prime case

- In theory, reduces constant in complexity of function field sieve.
  - Regardless of Kummer extension or not
  - Individual descent unchanged from [JL06]

- In practice, Kummer extensions very useful for records:
  - First 1175-bit field $\mathbb{F}_{p^{47}}$ with $p$ close to $2^{25}$
  - Then 1425-bit field $\mathbb{F}_{p^{57}}$ with $p$ close to $2^{25}$
  - Previous finite field record was 923 bits

- $47 = 6 \cdot 8 - 1$
- $57 = 7 \cdot 8 + 1$
Assume $p = 2$ to simplify exposition.

Define finite field by a relation:

$$x^{2^\ell} = \frac{h_0(x)}{h_1(x)},$$

which gives degree $k = \deg(I(x))$ extension, where $I(x)$ is a divisor of $h_1(x)x^{2^\ell} - h_0(x)$.

We have a systematic relation:

$$x^{2^\ell} - x = \prod_{\alpha \in \mathbb{F}_{2^\ell}} (x - \alpha).$$
Small characteristic – Basic idea [J13b]

- Use more general change of variable: \( x = \frac{aX + b}{cX + d} \), we get:

\[
(cX + d) \cdot (aX + b)^{2^\ell} - (aX + b) \cdot (cX + d)^{2^\ell} = \\
(cX + d) \cdot \prod_{\alpha \in \mathbb{F}_{2^\ell}} ((a - \alpha c)X + (b - \alpha d))
\]

- Moreover, after expanding the left-hand size, we find:

\[
(ca^q - ac^q)X^{q+1} + (da^q - bc^q)X^q + (cb^q - ad^q)X + (db^q - bd^q),
\]

where \( q = 2^\ell \).

It becomes a low degree polynomial after multiplying by \( h_1 \)
and replacing \( h_1(X) X^q \).

- As a consequence, multiplicative relations are very easy to find
Small characteristic – Resulting Complexity [J13b]

- Logarithms of smoothness basis in polynomials time
  - Because base field is very small compared to extension field
- Hard part is individual logarithms
  - Usual descent algorithm not good enough
  - Need to be completed by new descent algorithm (based on resolution of bilinear systems of Equations)
- Resulting complexity is:
  \[ L(1/4 + o(1)). \]

- Practical application:
  - New records in \( \mathbb{F}_{2^{4080}} \) and \( \mathbb{F}_{2^{6168}} \) recently announced
Descent phase

- In practice, bootstrap using continued fractions
- Classical descent (for high to mid degrees):
  - Consider $F(X, Y)$ of low degree in $X$ and $Y$; let $r \approx \ell/2$
  - We have:
    \[
    (F(X, X^{2r}))^{2^{\ell-r}} = F^*(X^{2^{\ell-r}}, \frac{h_0(X)}{h_1(X)}).
    \]
  - To apply descent to $f$, find $F$ such that $f|F(X, X^{2r})$
- New descent (for mid to low degrees):
  - Find $k_1$ and $k_2$ such that
    \[
    f|k_1^* \left( \frac{h_0(X)}{h_1(X)} \right) k_2(X) - k_1(X)k_2^* \left( \frac{h_0(X)}{h_1(X)} \right)
    \]
  - Gives relation between above polynomial and
    \[
    k_1(X) k_2(X) \prod_{\mu \in \mathbb{F}_{2^\ell}^*} (k_1(X) - \mu k_2(X)).
    \]
Descent phase — Complexity

- One step of classical descent from $D$ to $\mu D$
- $D$ linear conditions, i.e. $D$ monomials of degree $\sqrt{D}$ in each of $X$ and $Y$.
- Degrees on sides of equation:
  - Left side: $2^r \sqrt{D} - D \approx \sqrt{D}q$
  - Right side $(2^\ell - r + \max(\deg h_0, \deg h_1) \sqrt{D} \approx \sqrt{D}q$
  - Smoothness probability has $\log - \rho \log \rho$ with:
    $$\rho = \frac{2\sqrt{D}q}{\mu D}$$
- If $D \geq \sqrt{q \log q}$, prob is better than $L(1/4)$
Descent phase — Complexity

- One step of new descent from $D$ to $D - d$
- Bilinear system with $(D - d) + d$ vars
- Complexity of bilinear system is exponential in small number of vars:
  - We choose $d = O(q^{1/4} \log^{1/2} q)$
  - Thus top level dominates and complexity is
    
    $$\exp(O(q^{1/4} \log^{3/2}(q))).$$
  

Antoine Joux UVSQ and CryptoExperts
Revisiting discrete logarithms in medium/small characteristic
Descent phase — Breaking news

- New descent phase with Barbulescu, Gaudry, Thomé
- Without Gröbner bases
- Improved complexity:

  \[ \exp(O(\log q \log k)). \]

- Sub-exponential but not practical (yet)
Conclusion

Questions ?